DOCUMENT RESUME

ED 388 698 TM 024 089

AUTHOR Ferguson, William F.; Johnson, James T.
TITLE The NTE Shrinking-Score Phenomenon in Prediction of a CMEE-Mathematics Cut-Score, 1991-1993.
PUB DATE Nov 94
NOTE 14p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (Nashville, TN, November 9-11, 1994).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cutting Scores; Elementary Secondary Education; *Licensing Examinations (Professions); *Mathematics Tests; *Prediction; Teachers; *Test Results
IDENTIFIERS *Content Mastery Examinations for Educators; Mississippi; *NTE Core Battery; NTE Test of Professional Knowledge

ABSTRACT

A phenomenon previously noted in prior prediction of mathematics scores on the Content Mastery Examinations for Educators (CMEE) of the National Teacher Examinations Core Battery subtests of General Knowledge (NTE-GK) and Communication Skills (NTE-CS) was investigated. Prior research with 1991-1992 data sets had established an equation for the prediction of CMEE mathematics required cut-score of 340 for Mississippi teachers from the NTE scores. The subsequent addition of more data sets revealed that the cut-score could now be predicted less efficiently than before. Correlations between CMEE-Math and the NTE subtests now show far less relationship between them. The most reasonable explanation seems to be that the lower performing examinees on the NTE subtests have quit attempting the CMEE-Math, while examinees who typically do well on all standardized tests, but who lack the whole range of mathematics competencies, are still trying the CMEE-Math test. Coupled with the mathematically competent, who normally do reasonably well on the NTE subtests, the resulting pool of data sets appears to be random CMEE-Math scores. As a result, prediction is no longer possible. (Contains six tables and one figure.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 388 698

THE NTE SHRINKING-SCORE PHENOMENON IN PREDICTION OF A CMEE-MATHEMATICS CUT-SCORE, 1991-1993

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
WILLIAM F. FERGUSON
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC).

by

William F. Ferguson, Ed.D.
Educational Leadership and Research
University of Southern Mississippi
Hattiesburg, MS 39406-9221
(601) 266-4566

and

James T. Johnson, Ph.D.
Academic Computing Center
University of Southern Mississippi
Hattiesburg, MS 39406-5171
(601) 266-4103

A Paper
Presented at the Annual Meeting of the
Mid-South Educational Research Association
Nashville, TN
November 9-12, 1994

Running Head: CMEE-Math/MSERA94

TM024089

BEST COPY AVAILABLE

Abstract

The purpose of this study was to investigate a phenomenon noticed in prior prediction of Mathematics scores on the Content Mastery Examinations for Educators (CMEE) from the National Teacher Examinations Core Battery subtests of General Knowledge (NTE-GK) and Communication Skills (NTE-CS).

Prior research with 1991-1992 data-sets had established an equation for the prediction of the CMEE Mathematics (CMEE-Math) required "cut-score" of 340 for Mississippi teachers from the NTE scores. The subsequent addition of more data sets revealed that the cut-score could now be predicted less efficiently than before. Correlations between the CMEE-Math and the NTE subtests now show far less relationship between them.

The most reasonable explanation seems to be that the lower performing examinees on the NTE subtests have quit attempting the CMEE-Math, while examinees who typically do well on all standardized tests but who lack the whole range of mathematics competencies are still trying the CMEE-Math test. Coupled with the mathematically competent, who normally do reasonably well on the NTE subtests, the resulting pool of data-sets results in what appear to be random CMEE-Math scores. Voila! No prediction is now possible.

THE NTE SHRINKING-SCORE PHENOMENON IN PREDICTION OF A CMEE-MATHEMATICS CUT-SCORE, 1991-1993

The genesis of this study began with earlier studies of the relationships between and among the Content Mastery Examinations for Educators (CMEE) and the three National Teacher Examinations (NTE) Core Battery subtests of General Knowledge (NTE-GK), Communication Skills (NTE-CS), and Professional Knowledge (NTE-PK). In the first study, absolutely nothing was known of the CMEE for any subject areas, so emphasis was given to using preliminary data-sets to compute intercorrelational matrices and to predict the required "cut-scores" on the various CMEE tests from NTE subtests.

As any knowledgeable test expert might quickly point out, this could be considered a comparison of apples and oranges. All of the NTE tests are norm referenced, while all of the CMEE tests are criterion referenced. Be that as it may, they each result in scores which can be statistically compared. Consequently, statistical relationships can be determined that are of interest to researchers, to teacher educators, and to those who are responsible for determining appropriate scores for admission to teacher education programs and for certification purposes.

In 1990, 26 teacher education institutions across America agreed to participate in the validation study for the various CMEE tests. The University of Southern Mississippi (USM) was the only Mississippi institution which agreed to be a part of the validation process. Primarily, student teachers and education students at the

junior and senior levels participated. Of the ten USM persons who participated in the Mathematics test, the average percentage correct was .53, while for the composite group of mathematics majors in the total validation process nationwide the percentage was .55. Consequently, although no statistical test was performed, it was therefore assumed that the local mathematics population was not significantly different than the national population.

Following a a rigorous item analysis of the test items used in the validation study, the resulting items which were deemed acceptable were submitted to a panel of Mississippi mathematical "experts" selected to further examine the test items for clearness of expression, range of mathematics topics covered, difficulty level of items, and appropriateness in general for Mississippi mathematics teachers at the entry level through an "alternate route." This alternate route was developed primarily to allow non-education college graduates with mathematical expertise to qualify to become certified mathematics teachers in Mississippi. As such, the alternate route "cut-score" actually was set at a level said to be more rigorous than that for regular mathematics teachers.

Following the final Mississippi validation process, the CMEE was administered in Mississippi twice during 1991 and three times per year thereafter. At the first administration (July, 1991), seven of the 24 Mathematics examinees scored sufficiently high (340) to qualify for certification through the "alternative route." At the November, 1991 administration, however, only three of the 38 examinees scored 340 or better on exactly the same form of the CMEE tests for mathematics. During 1992, 10 of 43, six of 21, and 7 of 34 reached competence on that form. Finally, on the 1993 reports

for which data are available, nine of 29 reached competence in March and 11 of 35 in June. These and other useful data are provided below in Table 1.

Table 1: CMEE results for July, 1991 through June, 1993.

Date Administered	N	Passed	Mean	S.D.	Range
July, 1991	24	7	308.70	48.96	225-412*
Nov., 1991	39	3	291.23	43.91	228-439
March, 1992	43	10	303.56	48.96	235-412
June, 1992	21	6	313.95	44.50	241-403
Nov., 1992	34	7	304.06	44.09	210-398
March, 1993	29	9	310.31	40.67	235-387
June, 1993	35	11	309.74	54.51	232-473

*225 = Raw Score of 23 of 100 items
*309 = Raw Score of 53 (Mean of July, 1991 scores)
*340 = Raw Score of 65 ("Competence"--"Cut Score")
*412 = Raw Score of 86

As stated in an earlier paragraph, the average percentages of items correct for the USM and the national validation participants were 53% and 55% respectively, and this with volunteers. Is it unthinkable that the "volunteers" in each case were somewhat better than average at their sites? Isn't it typical that volunteers rarely represent the population from which they volunteered? And, doesn't research typically show that such volunteers (as a group) are somewhat above average, when compared to their population? At the moment, this is merely presented as "food for thought."

The first research presentation of CMEE-Math scores by this researcher was made in January of 1993, primarily using data obtained on USM examinees over 1991 and 1992. Complete data sets of CMEE-Math, NTE-GK, and NTE-CS were available on 44 subjects. Means and standard deviations were as follows:

Table 2: Means and standard deviations for NTE and CMEE scores, where N = 44.

	CMEE-Math	NTE-GK	NTE-CS
Mean	299.77	657.86	658.91
Standard Deviation	50.89	13.03	10.58

(N = 44 complete data sets)

Still later in 1993, new data were added, resulting in a grand total of 61 complete data sets. Now the statistics looked like this:

Table 3: Means and standard deviations for NTE and CMEE, where N=61.

	CMEE-Math	NTE-GK	NTE-CS
Mean	303.08	659.48	660.48
Standard Deviation	47.84	12.54	10.61

(N = 61 complete data sets)

Obviously, the added data sets had slightly higher scores in each respective position, since their addition to the original set of 44 data-sets led to increased means in all three cases.

In addition to the means and standard deviations, intercorrelational matrices are available for each of the above data-sets. They are provided below in a format that allows quick comparisons between the original group of 44 and the final group of 61, which resulted from the addition of 17 more complete data-sets over the next year or so.

Table 4: Intercorrelations of CMEE-Math and NTE Core Battery scores .

	NTE-GK N=44 / N=61	NTE-CS N=44 / N=61
CMEE-Math	.690 / .635	.435 / .399
NTE-GK		.712 / .695

Notice the reduced intercorrelational coefficient values which will lead directly to reduced efficiency in predicting CMEE-Math scores from the available NTE-GK and NTE-CS scores and from any hypothetical NTE-GK and NTE-CS scores in the future. For the N=44 data, the prediction of CMEE-Math by NTE-GK and NTE-CS yields a multiple R of .6946, a R^2 of .4824, and an adjusted R^2 of .4572. When enough additional data-sets were added to bring the total to 61, the respective values for multiple R, R^2, and the adjusted R^2 dropped to .6375, .4064, and .3859. Another interesting phenomenon appears when the prediction equation values are presented.

Prediction Equations:

(N=44): CMEE-Math = 3.0130 (NTE-GK) - 0.5513 (NTE-CS) - 1319.067

(N=61): CMEE-Math = 2.6393 (NTE-GK) - 0.3681 (NTE-CS) - 1194.292

Now watch what happens if the 61 scores are dichotomized in a different manner, now so that all 1991 scores are together (N=33) and all post-1991 scores are together (N=28). First of all, notice what happens to the means and standard deviations.

Table 5: Means and Standard Deviations for NTE and CMEE-Math when Dichotomously Grouped, N=33 and N=28.

	CMEE-Math N=33 / N=28	NTE-GK N=33 / N=28	NTE-CS N=33 / N=28
Mean	297.21 / 310.00	656.45 / 663.04	656.88 / 664.71
Stan. Dev.	51.19 / 43.45	13.971 / 9.67	10.87 / 8.71

When t-tests were performed on the above sets of means, all three post-1991 means were significantly higher than the 1991 means. With the increased mean values for the post-1991 scores, either the lower scoring examinees are now trying the CMEE-Math in greatly reduced numbers or else many more higher scoring examinees are now trying the CMEE-Math. With the greatly decreased standard deviation values, the latter explanation now seems more plausible. Furthermore, this conclusion was confirmed by visual examination of the data-sets.

Table 6: Intercorrelations of CMEE-Math and NTE Core Battery scores when dichotomously grouped, N=33 and N=28.

	NTE-GK N=33 / N=28	NTE-CS N=33 / N=28
CMEE-Math	.744 / .400	.576 / .037
NTE-GK		.792 / .405

While the correlational coefficients for the 1991 data-sets appear to be reasonably appropriate, the coefficients for the post-1991 data sets--if mathematically correct--absolutely destroy any possibility of finding a valid prediction equation using those data-sets. The 1991 data-sets yield multiple R, R^2, and adjusted R^2 values of .7743, .5540, and .5243 respectively. And, post-1991 data-sets yielded respective values of .4224, .1784, and .1127,

with a non-significant analysis of variance. The prediction equation for the 1991 data is as follows:

$$\text{CMEE-Math}_{(N=33)} = 2.8250\ (\text{NTE-GK}) - 0.1603\ (\text{NTE-CS}) - 1451.9658$$

The above equation still has the odd circumstance whereby increasing values of NTE-CS for any given value of NTE-GK will result in decreased values for CMEE-Math. However, the predictive efficiency of the above equation is superior to any previous prediction equation. That is, for the same values for NTE-GK and NTE-CS being "plugged into" any of the previously listed equations, the one immediately above predicts the higher CMEE-Math score. Most researchers would want to know why that would be true, since the addition of more data-sets to a relatively small data-set usually adds to the validity and reliability of any predictive ability of the data-sets. That definitely is not the case with these scores. Why?

The answer to the puzzle this time may not lie in additional high powered statistics; rather, it may lie in an "autopsy" of the individual data-sets as might be performed by an accountant attempting to locate an error in an accounting report. Obviously, on the first time the CMEE-Math test was administered, all examinees began equally--not knowing much about the test format, the content being tested, the difficulty-level, what the passing [competency] score was, or how it related to any other known score. Basically, they were all groping in the dark.

Initially, about all that potential examinees knew was that Mississippi had changed its original "alternate route" procedure from scoring at the 51st percentile on all four parts of the NTE.

The new scheme required them to score only at the 25th percentile on NTE-GK and NTE-CS and to "pass" the appropriate CMEE section, in this case mathematics. And just what does it mean to "pass" the CMEE-Math test? No one knew, and the only way to find out was to try it. The passing score couldn't be too tough; after all, hadn't the NTE score been lowered to the 25th percentile? Surely the CMEE-Math standard was also near the 25th percentile. Wrong; the required CMEE-Math score of 340 is roughly equivalent to a NTE-GK score of 674, or approximately at the 75th-80th percentile for the typical NTE examinee, including all fields. For just mathematics majors, it still will be slightly above the 50th percentile.

With absolutely no real conception of what they were getting into, hopeful mathematics teachers came to the first administration of the CMEE tests. With very little publicity for the first administration, seven of the 24 examinees reached the criterion level for passing. With more publicity, 38 showed for the second administration; and, only three passed. On the third administration, however, 10 of 43 passed; several of these were repeaters from a previous administration. And this is where the prediction equation begins to become muddled.

Contrary to what the researcher had concluded and stated in previous research and in the initial abstract submitted to the Mid-South Educational Research Association, it has been the mathematically weak student who learned quickly that the passing standard was very much out of their reach. By and large, this group has declined in numbers and in relative percentage over time. On the other hand, the mathematically knowledgeable and the traditionally good test-takers have persisted, often to reach the

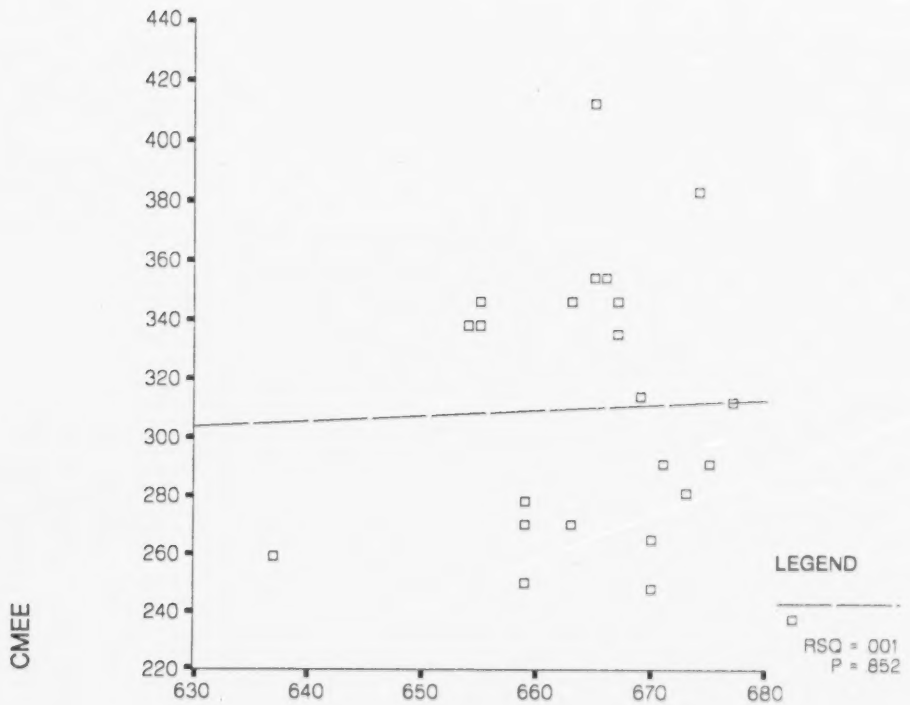
required score of 340. Some examinees participated in five of the first seven administrations, and many who scored "decently" initially returned several times before either achieving success or else determining that they couldn't make the score. Score patterns for a few examinees suggest that they soon concluded that they were being administered precisely the same form of the CMEE-Math test each time. Often there eventually came a quantum leap, and the required score was obtained.

Final thoughts: Assume there is an examinee whose "true" mathematical knowledge base would place him with a CMEE-Math score of 330. With unlimited repeated testing, even assuming no knowledge gained from prior test administrations, the law of averages dictates that eventually the person may make 340. Furthermore, a "test-wise" examinee who realizes that only one form of the test exists will, given enough opportunities, soon identify enough items to reach the required score. Does this examinee's score indicate "competence," or does it indicate primarily "test-wiseness"? Even with these potential threats to validity recognized, the current level of "competence" may very well be so high that only those with legitimate mathematical knowledge can ever reach it. But this does not consider the problem which will arise if and when this examination becomes accepted as a [or THE] common instrument for regular certification. If the current level was set at that point because it was supposed to represent a competent level of mathematical knowledge rather than adhere to some norm-referenced standard, what will happen when this instrument becomes the primary instrument for even "regular" mathematics certification? At that point, less mathematics

education majors than at present will be successful in passing the certification exam. But how can the current CMEE-Math score be lowered if the current score was set because it was deemed the minimumly desired level of competence? Can a lower score still be called "competence"?

These and other questions and concerns need to be addressed. Rather than be considered simply criticism of the CMEE-Math, this paper addresses findings and conclusions stemming from analyses of current CMEE-Math data. CMEE and Mississippi officials should address the issues raised here and attempt to rectify them in order to maintain and/or improve the current level of mathematics teachers.

1992


440
420
400
380
360
340
320
300
280
260
240
220
630
640
650
660
670
680
CMEE
LEGEND
RSQ = 001
P = 852
NTE COMMUNICATION SKILLS